DOT/FAA/AM-95/9

Office of Aviation Medicine
Washington, D.C. 20591

Functions of External Cues in Prospective Memory

O.U. Vortac
Mark B. Edwards
University of Oklahoma
Norman, Oklahoma 73019

Carol A. Manning
Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma 73125

DTIC
ELECTE
MAR 1 6 1995
S G D

February 1995

Final Report

19950313 043

This document is available to the public
through the National Technical Information
Service, Springfield, Virginia 22161.

U.S. Department
of Transportation
Federal Aviation
Administration

Technical Report Documentatio

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-95/9		

4. Title and Subtitle	5. Report Date
Automation and Cognition in Air Traffic Control: An Empirical Investigation	February 1995
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
O.U. Vortac, Mark B. Edwards, Dana K. Fuller, and Carol A. Manning	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
FAA Civil Aeromedical Institute P.O. Box 25082 Oklahoma City, OK 73125 University of Oklahoma Dept. of Psychology 755 W. Lindsey, Room 705 Norman, OK 73019-0535	11. Contract or Grant No. DTFA-02-91-C-91089
12. Sponsoring Agency name and Address	13. Type of Report and Period Covered
FAA Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, SW. Washington, DC 20591 FAA Research and Development Service Federal Aviation Administration 800 Independence Avenue, S.W. Washington, DC 20591	14. Sponsoring Agency Code

15. Supplemental Notes

This work was performed under task AAM-B-94-HRR-141 and Contract No. DTFA-02-91-C-91089.

16. Abstract

A simulation of an air traffic control task was the setting for an investigation of the functions of external cues in prospective memory. External cues can support the triggering of an action or memory for the content of the action. We focused on memory for the content and manipulated the temporal characteristics of the external cue to disentangle two possible functions the cue can support: 1) An external cue visible during a retention interval could support rehearsal of the to-be-performed action; 2) An external cue visible at the end of a retention interval could support retrieval of the to-be-performed action. Two experiments were conducted that converge on the same conclusion: the primary function of an external cue was to support retrieval. Implications for the design of a computer interface to present prospective cues are discussed

17. Key Words	18. Distribution Statement
Automation, Air Traffic Control, Flight Progress Data, Cognitive Psychology, Memory, Applied Psychology	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	17	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Acknowledgments

O. U. Vortac represents the collaborative efforts of Francis T. Durso, Scott D. Gronlund, and Stephan Lewandowsky, all of the Department of Psychology, University of Oklahoma, Norman, OK 73019 (Email: sgronlund@uoknor.edu).

This work was supported by Contract DTFA-02-91-C-91089. Thanks to Bob Blanchard, Bill Collins, Doug Herrmann, Paula Hertel, Mark Rodgers, Dave Schroeder, Earl Stein, Mike Wayda, and especially the reviewers and action editor, for edifying comments on earlier versions of this manuscript. We express our appreciation to Dusty Godwin, Allen Molloy, Leslie Wiggins, and Russ Brown, for their fine efforts on this project.

FUNCTIONS OF EXTERNAL CUES IN PROSPECTIVE MEMORY

The sense of memory commonly evoked and the sense that has dominated research in psychology for over a century is retrospective memory, or memory for past events. The use of retrospective memory is usually prefaced by the remark, "Remember when...." However, memory is also often used to remember to do things in the future. This can range from remembering to turn off the iron, remembering to pick-up grandma at the airport, to remembering to take medication. This kind of memory has been termed prospective memory by Baddeley and Wilkins (1984) and Meacham and Leiman (1982). The use of prospective memory is usually prefaced by the remark, "Don't forget to...."

Following Einstein and McDaniel (1990), we divide prospective memory into two components. One component is the *trigger* that initiates the prospective action; the other is the *content* of the action that is triggered. For example, if the task is to pass along a phone message to a colleague, seeing the colleague is the trigger; the content is the relayed message. Both are required for prospective memory to be successful. Forgetting to stop at the grocery store on the way home but remembering what was to be purchased is unsatisfactory. Likewise, remembering that you have a message for a colleague but not remembering the message content is unsatisfactory.

In prospective memory, external cues are pervasive (for a review see Harris, 1984). They are even commercially available, including a pill alarm, a whistling teapot, or a headlight alert, to name a few (see Petro, Herrmann, Burrows, & Moore, 1991). External cues support successful performance of prospective actions in one of two ways.

An external cue can support primarily the trigger component of prospective memory (a Post-It™ note by the door as a reminder to return an overdue book to the library, setting the alarm on your watch or computer, a string around the finger). A trigger cue proclaims that now is the time to take an action, but it does not indicate what that action is. For example, Meacham and Leiman (1982; Meacham & Singer, 1977) had subjects return postcards at various dates in the future. Half the subjects were given colored tags to put on their key chains as a reminder. Subjects with tags (external cues) did better than those without tags. The tags signaled that something was to be done, but did not indicate what that something was.

A second way that external cues can support prospective memory is by combining the trigger and content components (placing running shoes by the door to remember to take them to work, taping a note on the computer screen with a reminder to pick up bread on the way home). A combined cue proclaims that it is time to take the action as well as what the action should be. Meacham and Columbo (1980) instructed children to remind the experimenter to open a *surprise box* when the experiment was finished. The box was placed out of sight and half the children were given a clown to use as a reminder, and instructed how to use it. The clown served as a trigger because it was visible at the conclusion of the retention interval (the end of the experiment). The clown also served to remind the children of the content (to view the surprise box). The external cue was effective; more of the children who got the clown as a reminder asked to see the surprise box.

The Meacham and Columbo (1980) experiment illustrates the problem that motivates the current research. Because the clown was present throughout the retention interval, subjects could remember to ask for the surprise box for either of two reasons: 1) The clown could support rehearsal of the content during the retention interval (i.e., rehearsal of asking to see the surprise box later); 2) the clown could support retrieval of the content (i.e., retrieval of asking for the surprise box) at the end of the retention interval. These two functions are naturally confounded if the external cue is present throughout a retention interval. We want to know to what extent successful

performance is due to the subject rehearsing the content during the retention interval, versus seeing the cue and retrieving the content when they can execute it.

The primary goal of the present research is to examine how external cues support memory for the *content* of a to-be-performed action. We disentangled the rehearsal from the retrieval function of a combined cue by manipulating its temporal availability. However, due to this manipulation, the combined cue was not an equally effective trigger in all conditions. Therefore, we had to add a second trigger cue that indicated nothing about content. One cue supported content only, the other supported the trigger only.

Before describing in detail the external cues that were used, we describe the setting for the experiment. We felt it was important to maintain the ecologically-valid approach characteristic of much of prospective memory research (e.g., Meacham & Leiman, 1982; Meacham & Singer, 1977; West, 1988; Wilkins & Baddeley, 1978), although we required a degree of control characteristic of a laboratory-based task (e.g., Einstein & McDaniel, 1990; McDaniel & Einstein, 1993). The following task was chosen.

The Federal Aviation Administration's Air Traffic Scenario Test (Aerospace Sciences Inc., 1991) is a PC-based simulation that provides a fairly complex and realistic human factors setting, which we modified as needed. With one exception (Drew, 1940), work on prospective memory has not been conducted in a human factors setting. It is a timely period to examine prospective memory in air traffic control. Impending automation will soon change the way in which controllers interact with their external memory cues. Over the next few years, the human-computer interface in the en route[1] air traffic control environment will undergo substantial modifications, creating an opportunity to determine the optimal way in which prospective cuing should be accomplished.

Air traffic control

Air traffic control is a very rich domain in which to investigate prospective memory. Controllers make extensive use of prospective memory; they frequently have to put something off temporarily, either because the situation does not allow it or their workload is too great. When a controller cannot execute a control action immediately and has to remember to take that action in the near future, he or she will often make use of an external memory aid. This might happen if a pilot has requested the use of a higher altitude that cannot be granted until the aircraft is clear of crossing traffic.

The major ramification of the impending automation on the controller's work environment concerns the display of flight data not presented on the radar display. Presently, controllers receive these flight data on rectangular paper strips, called flight progress strips. These are the controllers' external memory cues. There is one strip for each controlled aircraft on the radar display, each with 31 fields of information, including the call sign, planned altitude, and destination. The controller has various legal responsibilities that require extensive interaction with and marking of the strips. Hopkin (1988) and Jackson (1989), among others, have speculated that, although this marking of the strips is at least partially done for legal reasons, it likely serves to support many cognitive functions that aid controllers in maintaining their "picture" of the airspace (for a review see Cox, 1992; Vortac & Gettys, 1990).

Prospective memory is one of these cognitive functions. In the current system, controllers mark directly on the paper flight progress strips, and will often offset one strip from the others as a prospective cue. The offset strip serves as the trigger; the information on the strip contains the required content regarding what action should be taken. For example, a pilot may request a lower altitude, but the controller cannot do so until the plane clears crossing traffic. The controller may offset that plane's strip as a future reminder to take this action, and write the requested altitude on the strip to help remember the content of the future action.

Under the planned automation, these paper strips will be replaced by electronic strips on a computer screen. In the new system, controllers will interact with electronic strips indirectly via a keyboard and

[1] En route controllers are responsible for aircraft traveling between terminal facilities across the nation.

trackball. In addition, many updating functions will be taken over by the computer. Although the highlighting of an electronic strip as a prospective reminder will be possible, it may be less effective than the controller reaching up and offsetting the strip as is done in the current system; highlighting would be initiated through the keyboard, like so many other routine actions. In the retrospective memory literature, it is well-known that memory is better for something you do as opposed to something done for you (e.g., Slamecka & Graf, 1978); the same may be true in prospective memory. There are reasons for concern regarding automation, as well as opportunities to make recommendations to improve the design of the new system.

Vortac, Edwards, Fuller, and Manning (1993) compared prospective memory performance in two groups of controllers.[2] The *Normal* group controlled traffic as they normally would. The *Automation-Analog* group was forbidden from writing or manipulating the paper strips; this group could only look at the strips. The idea was that restricting interaction to looking only would simulate the lost functionality resulting from automation. At three different times during a scenario a prospective request was made. A pilot requested an altitude change or rerouting a couple minutes prior to entering the controller's airspace. Because controllers cannot unilaterally assume control of an aircraft outside of their assigned airspace, this created a natural prospective memory request. The Automation-Analog controllers granted more of the prospective requests, and did so sooner, than the Normal controllers. It is important to capitalize on this potential benefit of "automation" to prospective memory by considering how the new system should present external cues to the controller.

Experimental design

The four ways that the availability of content will be manipulated are: Both, Retrieval-only, Rehearsal-only, and Neither.

Rehearsal and Retrieval (Both). If the external cue is visible throughout the retention interval, both rehearsal and retrieval are supported. If the goal is to return a library book by the end of the day, the Both condition is akin to placing the book by the office door where it is visible throughout the day (the retention interval) and when you leave (the opportunity to execute the command). In the air traffic control situation, the controller would encode a prospective command that would remain visible throughout the retention interval. This electronic strip would support rehearsal and retrieval because the controller would be free to review/rehearse the information at any time, and the information would be available at the end of the retention interval.

Retrieval-only. Alternatively, the book could be placed by the office door, but then your day is spent in the lab, returning to the office just before heading home. The book would not support rehearsal during the day, but would be visible when you leave. The controller could encode the prospective command with an attached time delay (*change the altitude on American 123 in 2 mins*). Nothing would happen on the electronic strip until the end of this time delay when the prospective command would reappear. Although the command could be rehearsed during the retention interval, rehearsal would not be supported by the external cue.

Rehearsal-only. The book could be visible throughout the day but not when you leave. This would happen if you spent most of the day in the office but depart for home from the lab. The controller would encode the prospective command, which would remain visible on the electronic strip until sometime prior to when the command was to be executed. The strip would support rehearsal, but it would not support retrieval because it was not visible at that time.

Neither. Finally, the book could be placed by the office door but then the day is spent in the lab and you depart from there. The controller would encode the prospective command, but it would never appear again and the controller would rely solely on internal memory aids. The electronic strip would support neither rehearsal nor retrieval in this case.

Hypotheses. Intons-Peterson and Fournier (1986) found that external cues supported both rehearsal and retrieval functions in retrospective memory. For ex-

[2]In contrast to the present experiments, ATC instructors were used by Vortac, et al. (1993), and the simulation facility used was the high-fidelity simulator at the FAA's Mike Monroney Aeronautical Center.

ample, recall of a grocery list was better if subjects made notes (an external cue) to help them remember, even if the notes were not consulted at retrieval. Recall was higher still when the notes were consulted. Is the same true of prospective memory? If so, the Both condition should set the upper bound on performance. The Neither condition should set the lower bound on performance; subjects must rely solely on internal aids.

If Retrieval-only performance is equivalent to Both, and Rehearsal-only is equivalent to Neither, the external cue must support primarily a retrieval function. If Rehearsal-only performance is equivalent to Both, and Retrieval-only is equivalent to Neither, the external cue must support primarily a rehearsal function. Alternatively, Retrieval-only and Rehearsal-only could fall in-between the upper and lower bounds, implying that an external cue supports both functions.

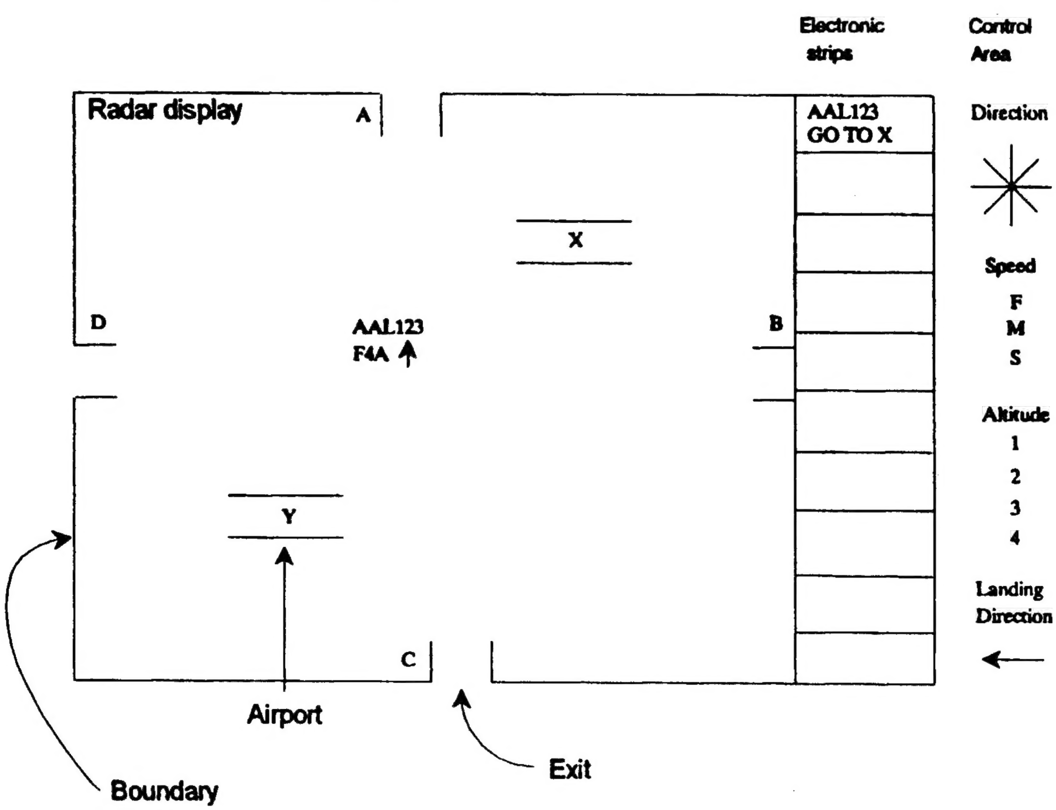
In Experiment 1, Both, Retrieval-only, Rehearsal-only, and Neither were manipulated within-subjects. Because subjects did not know how long the command would remain visible during the retention interval, or whether it would reappear, the use of different rehearsal strategies should be minimized. Controlling rehearsal and strategy in this way helped isolate availability of content as a causal factor.

Experiment 1

Method

Subjects. Subjects were recruited from the University of Oklahoma campus community and paid $5 an hour for their participation. Thirteen subjects participated for 12 days; eight participated for 6 days (after having first completed 4 days in Experiment 2). Those subjects for whom Experiment 1 was actually their

Figure 1. The airspace, including four gates (A, B, C, and D), two airports (X and Y), the column of strips including call sign and prospective command, and the control area for changing direction, speed, and altitude.



second experiment performed better overall than the other subjects, but the pattern of results was the same for both groups.

The air traffic control simulation. The experiment was conducted on microcomputers that ran the FAA's Air Traffic Scenario Test (Aerospace Sciences Inc., 1991). This simulation was a modified version of a test used by the FAA to screen applicants for air traffic controller positions. The simulation captures many of the major activities included in air traffic control and provided a rich environment in which to explore prospective memory. Figure 1 illustrates the screen display of the simulation, which consists of a radar display or airspace (far left), and an aircraft control area (far right). The modification we added to the original software was the column of electronic strips between the airspace and the aircraft control area.

The airspace consisted of four exits out of the airspace (labeled *A, B, C,* and *D*) and two airports (*X* and *Y*). Aircraft were represented in the airspace by a call sign (an alphanumeric identifier like AAL123), a directional arrow, and a data block. The data block, which moved as the aircraft moved, displayed the aircraft's speed (Fast, Medium, or Slow), altitude (1-low to 4-high), and destination (airport or exit). Aircraft AAL123 is illustrated in Figure 1, flying at Fast speed, altitude 4, towards its original destination at A.

The electronic strip for AAL123 is also illustrated. The second line of a strip is where the prospective command was displayed; it shows a change to the original destination (GO TO X). The original destination remained visible on the data block; the electronic strip was the only place to get information regarding the changed destination.

Procedure. Subjects began by observing the Experimenter control a five-minute practice scenario. The subjects then completed the practice scenario themselves. Subjects completed two 22-minute scenarios per day. The assignment of scenario to subject was counterbalanced, and no subject ever completed the same scenario twice.

The rules of the simulation were: 1) aircraft must exit a gate high and fast, 2) aircraft must land at an airport low and slow, and in the correct direction (the arrow at the bottom right-hand corner of Figure 1 indicated the landing direction), 3) aircraft could not overfly airports, but could overfly each other if they were at different altitudes, 4) aircraft had to remain separated from each other and the airspace boundaries, and 5) subjects should be as efficient as possible (i.e., get an aircraft to its destination as quickly as possible).

Interaction with the simulation was accomplished with a mouse. When a change to an aircraft's direction of flight, speed, or altitude was desired, the mouse was moved over the aircraft and clicked, thereby selecting the aircraft. Then the mouse was moved to the aircraft control area and a new direction of flight, speed, or altitude was selected.

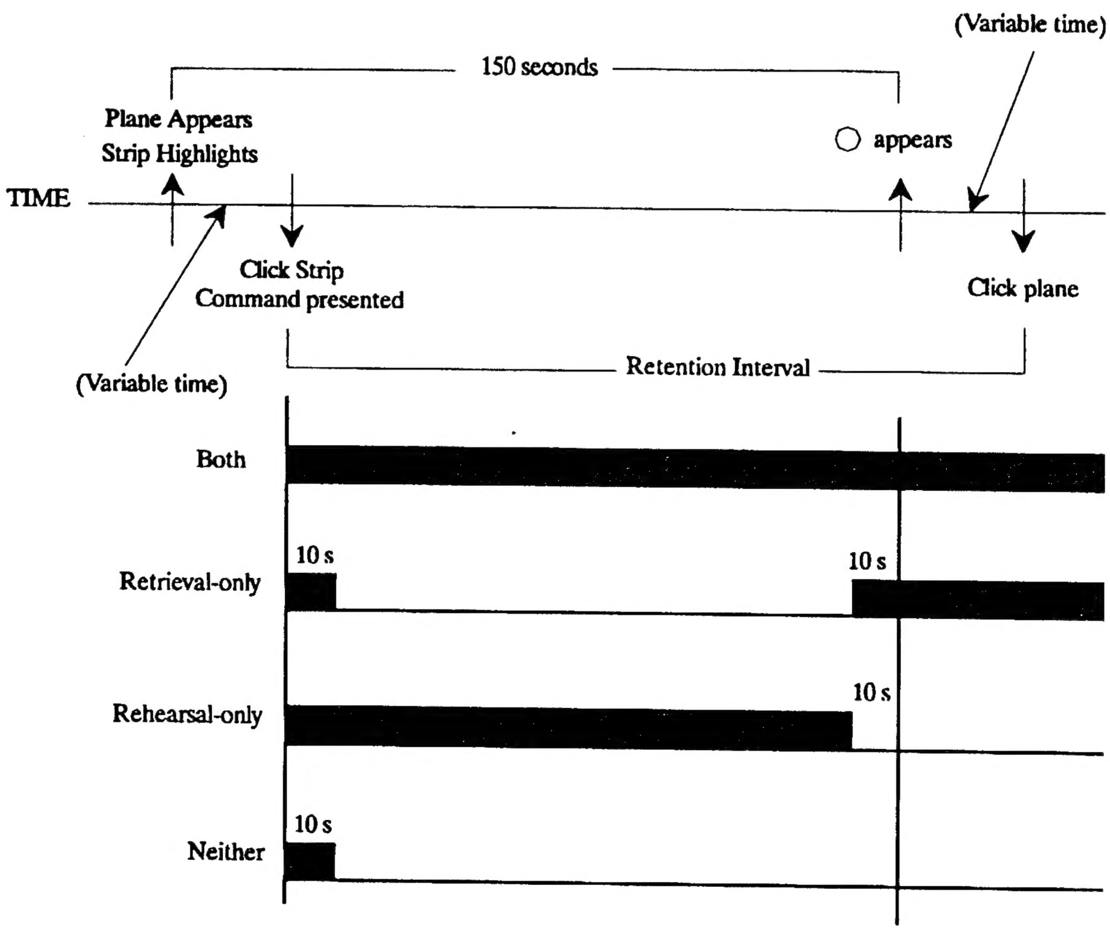
Figure 2 is a timeline of actions that illustrates how the prospective command was presented. Whenever a new aircraft appeared on the screen, the accompanying electronic strip appeared at the same time, highlighted in blue, with only the call sign visible. The subject was instructed to click the strip as soon as possible, whereupon the strip would return to its normal white on black and reveal a prospective command (if the aircraft were a critical plane). Note that the 150 s that elapsed between when an aircraft appeared with the strip highlighted to the time that the plane was able to fly, was for the subset of aircraft related to the prospective manipulation. (Subjects had been told during the instructions that the delay was to allow threatening weather to clear.) After 150 s, a small circle (O) appeared next to the data block (as did a non zero altitude). A click of the O initiated aircraft movement. Nothing on the strip could serve as a trigger, only the O next to the data block.

For the Both, Retrieval-only, Rehearsal-only, and Neither conditions, clicking the strip revealed a prospective command in the form of a changed destination. The prospective command was visible for the next 10 s. In the Neither condition, the command

Figure 2. Illustrates the temporal availability of the prospective command for the Both, Retrieval-only, Rehearsal-only, and Neither conditions. The timeline at the top of the figure indicates the sequence of actions to be taken. Beginning at time 0 on the far left: A plane appears, together with the appearance of its electronic strip (highlighted in blue). After some variable time, the subject clicks the strip to reveal the prospective command. 150 s after the plane appears, ❍ appears. A click of the ❍ initiates movement of the plane. The retention interval is measured from the time that the strip is clicked (revealing the command) until the ❍ is clicked; the command can now be executed.

The timeline for the delayed-no change condition was similar except that when the subject clicked the strip, no command appeared. The nondelayed-change condition was just like the Both condition except that there was no delay between when the plane and the ❍ appeared. That is, the plane was ready to go when it appeared. Finally, the nondelayed-no change condition was similar to the nondelayed-change condition except that there was no destination change.

The horizontal bars at the bottom two-thirds of the figure indicate the temporal availability of the prospective command for the four conditions of interest (the filled portion indicates when the command was present).



then disappeared, never to return. In the Both condition, the command remained visible until the aircraft left the airspace. In the Rehearsal-only condition, the command disappeared 10 s before the O appeared. In the Retrieval-only condition, the command reappeared 10 s before the O appeared. The content component of prospective memory was measured by the accuracy of sending these aircraft to their correct changed destination.

As is evident from Figure 2, the subject's retention interval could be longer or shorter than 150 s depending on when the subject responded to: 1) the highlighted strip, and 2) the appearance of the O. For example, if the subject was slow to respond to the highlighted strip (e.g., it took 30 s) and quick to respond to the O (e.g., 5 s), the retention interval (the time from the initial encoding of the prospective command to the time that the aircraft was moving and the prospective command could be executed) was 125 s. On the other hand, if the subject was quick to respond to the highlighted strip (5 s) and slow to respond to the O (30 s), the retention interval was 175 s. The retention interval was under subject control and was a function of a variety of factors, most notably workload.

The trigger component of prospective memory was measured by the latency to click the O. Air traffic control is an event-based task (see Einstein & McDaniel, 1990). A trigger signals that an action can be taken, but does not require that the action be taken right now. There was no specific time at which the prospective command was to be executed (as long as O was present on the relevant aircraft). Subjects were simply instructed to be as efficient as possible, which meant that they should try to get to it as soon as possible. In contrast, a time-based prospective task requires O that an action must be performed at a specific time (e.g., taking cupcakes out of the oven, Ceci & Bronfenbrenner, 1985).

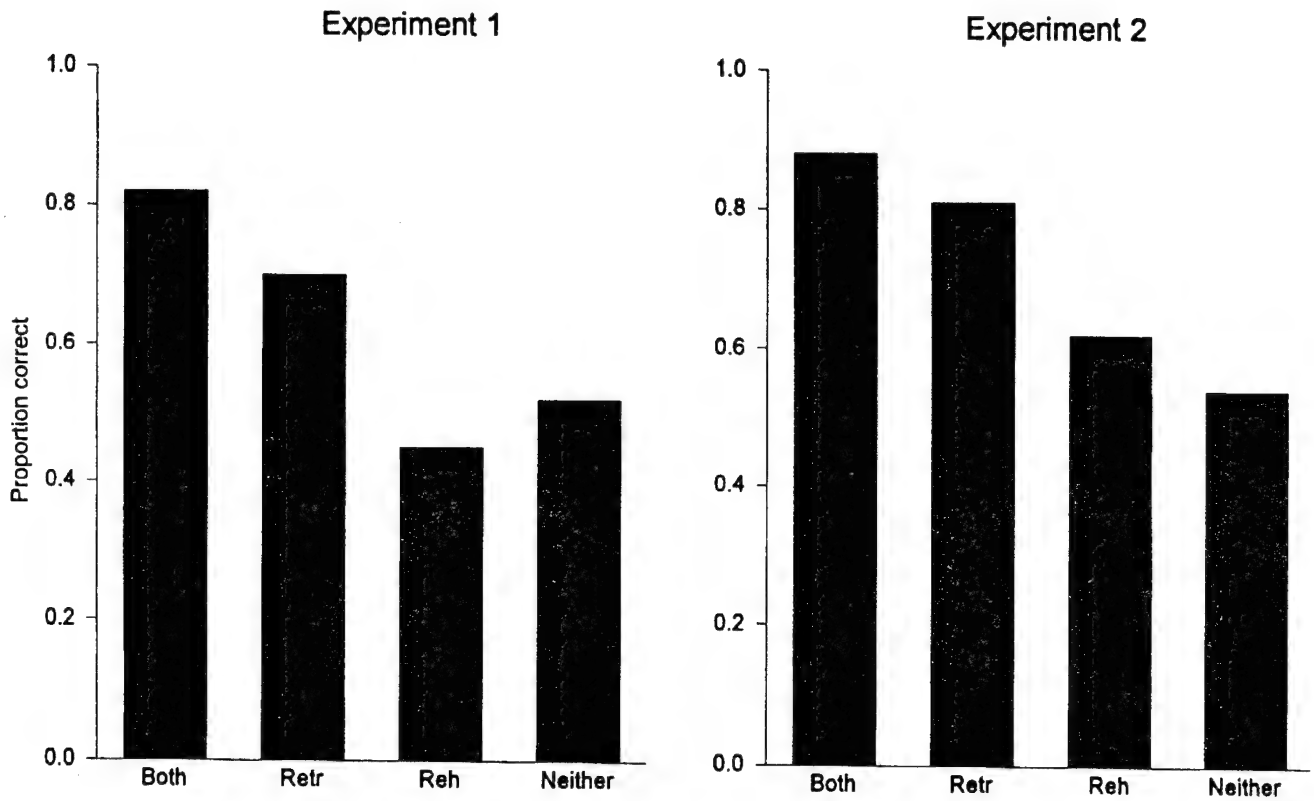
Design. In each scenario, there were 25 aircraft. The scenario began with three aircraft that were already flying, new aircraft appeared every 45 s thereafter. Only four of the 25 aircraft were the critical planes (those that were *delayed* and had their destination *changed*), one each for Both Retrieval-only, Rehearsal-only, and Neither. The ordering of these conditions was rotated through the 6th, 11th, 16th, and 21st aircraft that appeared in a scenario. The critical aircraft were spaced 3.75 minutes apart, so subjects had one on its way before the next appeared. Nevertheless, the load on prospective memory increased over the course of the scenario because it took an average of 13.3 minutes to get an aircraft to its destination.

In addition to these four critical aircraft, there were two aircraft that were *delayed* and their destination was *not changed*, and two that had their destination *changed* and were *not delayed*. They were randomly assigned to the 5th, 10th, 15th, and 20th positions in the scenario. The remaining 17 aircraft were *not delayed* and *not changed*. The non-critical aircraft served to disguise the true purpose of the experiment. For all aircraft, the O signaled that the aircraft could now fly. For the nondelayed aircraft, the O appeared when the strip highlighted; for the delayed aircraft, the O appeared 150 s after the strip highlighted. Also, not all the delayed planes required a destination change, only the four critical aircraft.

Scenario construction. The base scenario was constructed as follows: A set of X, Y coordinates was selected for each aircraft that specified its starting position. Next, we selected a direction of flight (8 possible), initial speed (always medium), original destination (16 went to exits, 4 to each, and 9 went to airports, 5 to one and 4 to the other), call sign (four each of SWA, UAL, DAL, TWA, NWA, and five AAL) with a three-digit number selected from a random number table, and an initial altitude (either 2 or 3). Twenty-four scenarios were constructed altogether.[3]

[3] The remaining 23 scenarios were constructed from the base scenario in the following way: First, the X, Y coordinates were rotated down (the starting position for aircraft 1 became the starting position for aircraft 2, and so on), the destination field was rotated up, a new random 3-digit number completed the call sign, and initial flight level was changed from 2 to 3 or vice versa. Half of the scenarios had the aircraft land from right to left, and half from left to right. Although it is difficult to intuit the result of these changes, the purpose of this procedure was simple—to create 24 unique scenarios.

Figure 3. Proportion sent to the correct destination. Experiment 1 data are given in the left-hand panel, Experiment 2 data in the right-hand panel. The Retrieval-only condition is abbreviated Retr and the Rehearsal-only condition is abbreviated Reh.



Results and Discussion

On the first day, subjects were interrupted frequently by the experimenter and reminded about how to proceed. Therefore, we counted the first day as practice and excluded it from the analysis. We examined data only for those retention intervals between 60 to 240 s, which accounted for better than 93% of the data.[4] One subject was excluded because more than 30% of aircraft failed to get to a destination.

The left-hand panel of Figure 3 gives the proportion of aircraft sent to the Correct destination for the Both, Retrieval-only, Rehearsal-only, and Neither conditions. Performance in all conditions was significantly above chance (or 1/6, smallest t (19) > 5.50). A one-way repeated-measures ANOVA detected a significant difference among conditions, $F(3, 72) = 7.00$, $MS_e = .07$, $p < .05$. Pairwise t-tests were conducted controlling α at .05/number of comparisons. They showed that the external cue supported primarily a retrieval function: the Both and Retrieval-only conditions did not differ, and both were better than the Neither condition. The Both condition was also better than Rehearsal-only condition; the Retrieval-only condition was marginally better than the Rehearsal-only condition (t (19) = 2.62, $p < .02$). The Rehearsal-only and Neither conditions did not differ.

The proportion that crashed and failed to reach any destination, or were still flying at the conclusion of the scenario both resulted in missing data. We combined these categories (called StillCrash).The StillCrash category occurred rarely ($M < 4.6\%$) and did not vary with condition, $F(6, 135) < 1.0$, $MS_e = .004$.

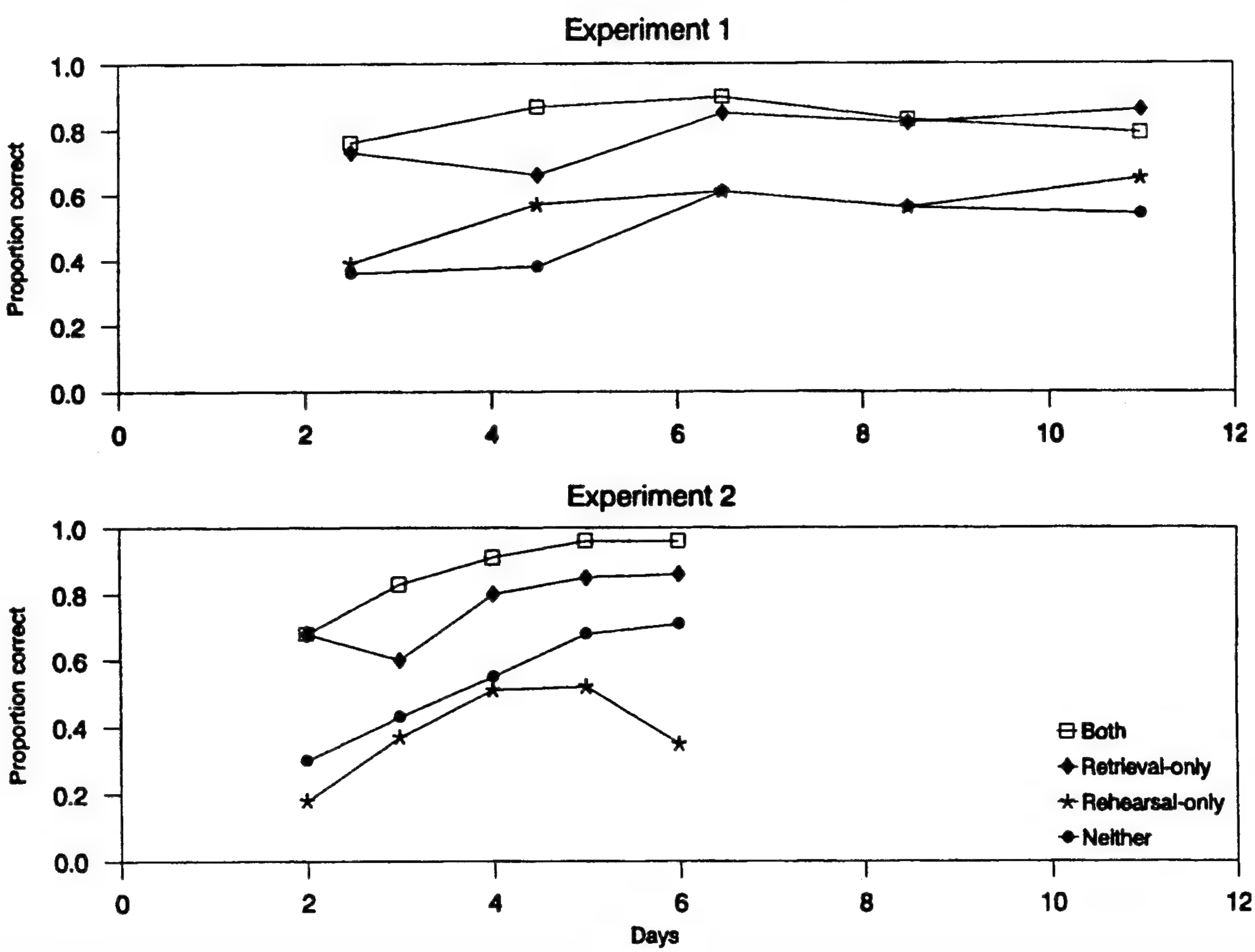
[4] In the Retrieval-only condition, we also excluded occurrences where the latency to click the highlighted strip was greater than 90 s. This ensured at least 60 s (i.e., the minimum retention interval) during which no command was visible. It was not necessary to do this in the other conditions because the command did not disappear and later reappear.

Most of the errors involved sending aircraft to the original destinations on the data block. That is, if subjects made an error, it was more likely to involve forgetting that a prospective command was issued for an aircraft (and sending it to the data block destination) than remembering that a command had been issued but not remembering its content (and guessing randomly among the five non-data-block destinations). However, this may have more to do with the fact that the data block destination was the correct destination 76% of the time.

The top panel of Figure 4 shows the proportion sent to the correct destination over days. The data were collapsed over adjacent days to smooth the function. Because Day 1 was excluded, Day 12 was combined with Days 10 and 11. Performance improved in all conditions through Days 6-7. Beyond that point performance was fairly stable. The learning data show that the relative contributions of rehearsal versus retrieval were unchanged by repetition over days and increasing familiarity with the task.

The second dependent measure involved the trigger: the time to click the O to get an aircraft flying. Although no one ever forgot to do this, on occasion it took over 100 s (means ranged from 17 to 28 s across conditions). To minimize the impact of the extreme outliers, we conducted an ANOVA on the geometric mean for each subject. We included both critical (the

Figure 4. Proportion sent to the correct destination by days. The B represents the Both condition, R the Retrieval-only condition, r the Rehearsal-only condition, and N the Neither condition. Experiment 1 data are given in the top panel, Experiment 2 data in the bottom panel.



four delayed-change conditions—Both, Retrieval-only, Rehearsal-only, and Neither) and non-critical aircraft (delayed-no change, nondelayed-change, and nondelayed-no change). There was no significant difference among conditions, $F(6, 135) = 1.96, p > .05, MS_e = 317.7$. The time to respond to the trigger was not affected by manipulations of the availability of content. An indirect measure of workload was also computed by determining how quickly a subject clicked a highlighted strip. Again, geometric means were used, and again, no significant difference was found among conditions, $F(6, 135) < 1, p > .1, MS_e = 283.6$.

In sum, Experiment 1 showed that an external cue supported a retrieval function. Making the prospective command visible during the retention interval was no better than never making it visible after an initial encoding period of 10 s. Will the same be true when rehearsal and retrieval strategies are unencumbered?

In Experiment 2, subjects were allowed to concentrate on a single method for presenting prospective commands. That way, differential rehearsal and retrieval strategies could be used in conjunction with the method of cue presentation to optimize performance. Perhaps subjects can compensate in the Rehearsal-only and Neither conditions by using a more elaborative rehearsal strategy. The cognitive cost of this may have been too great in Experiment 1, or the benefit too small; more elaborate strategies might be worth the effort in Experiment 2.

EXPERIMENT 2

The only change from Experiment 1 was that the Both, Retrieval-only, Rehearsal-only, and Neither conditions were manipulated between-subjects. This allowed subjects to optimize rehearsal and retrieval strategies.

Method

Subjects. Subjects were recruited from the University of Oklahoma campus community and were paid $5 an hour, or received experimental credit toward satisfying a course requirement, in exchange for their participation. Forty-four subjects participated for 6 days; 13 participated for only 4 days (due to a scheduling mix-up). Three subjects were excluded because over 30% of aircraft never reached a destination.

Procedure. Both, Retrieval-only, Rehearsal-only, and Neither were manipulated between subjects; subjects were randomly assigned to one of the four conditions ($N_{\text{Both}} = 14$, $N_{\text{Retrieval-only}} = 12$, $N_{\text{Rehearsal-only}} = 14$, $N_{\text{Neither}} = 14$). We selected 12 of the Experiment 1 scenarios to be used.

Results and Discussion

We counted the first day as practice and excluded it from the analysis. We examined data only for those retention intervals between 60 to 240 s (better than 94% of the total).[4]

The right-hand panel of Figure 3 gives the proportion of aircraft sent to the Correct destination. Performance in every condition was above chance (1/6, smallest t (13) > 3.0). A one-way ANOVA on the Both, Retrieval-only, Rehearsal-only, and Neither conditions showed a significant effect of condition ($F(3, 50) = 6.56, MS_e = .38, p < .05$). (The delayed-no change, nondelayed-change, and nondelayed-no change conditions were excluded because they were manipulated within-subjects.) The Ryan multiple-range test controlling α at .05/number of comparisons showed that the Both condition was better than Rehearsal-only and Neither, and Retrieval-only was better than Rehearsal-only. The Both and Retrieval-only conditions did not differ, nor did the Rehearsal-only and Neither conditions. The StillCrash category occurred rarely (M < 7%) and did not vary across conditions (F < 1).

The bottom panel of Figure 4 shows the Correct category over days. Day 1 was excluded. Performance improved through day 5; beyond that it appears that Both and Retrieval-only performance was stabilizing, although stabilization of the Both data may be a ceiling effect. Note that the means for Days 5 and 6 represent a subset of subjects comprising the means for the previous days. Performance in the Neither condition seemed to be improving with practice, while performance in the Rehearsal-only condition seemed to be worsening (although the latter was based on only five subjects). Perhaps subjects sensed that the

Neither condition was difficult and tried harder to succeed. The Rehearsal-only subjects, on the other hand, might have become increasingly frustrated at their inability to perform better at a task they may have perceived as relatively easy.

An ANOVA on the geometric mean time to click O was not significant ($F(3, 55) < 1$). (We excluded the delayed-no change, nondelayed-change, and nondelayed-no change conditions because they were manipulated within-subjects.) The time to respond to the trigger was not affected by manipulations of the availability of content. The ANOVA on the geometric mean time to click the highlighted strip was not significant ($F(3, 55) = 1.52$). Either time to click the highlighted strip was not a sensitive measure of workload, or subjects had no more trouble with the difficult conditions (e.g., Neither) than with the easier conditions (e.g., Both).

The findings of Experiment 2 parallel those of Experiment 1; the external cue supports primarily a retrieval function. Apparently, allowing the opportunity to tailor rehearsal and retrieval strategies to the way the external cue was presented made little difference. The retrieval function of the external cue was primary and was not affected by differential rehearsal and retrieval strategies.

General Discussion

Two experiments demonstrated that external cues support a retrieval function. In a simulated air traffic control task, it is important that an external cue that provides content be visible at the time that the prospective command can be executed. Having the external cue visible during the retention interval was no more effective than not having the cue visible at all. The benefit to the Both and Retrieval-only conditions was visible from the outset of training, and remained over the course of training. We also found that these manipulations of content did not affect the latency of an action being triggered, although this may have something to do with this being an event-based task.

When subjects could tailor their encoding and retrieval strategies to the cues (Experiment 2), there was a potential ceiling effect for the Both condition over the last three days of training. If this was a real ceiling effect, the Both condition would perhaps score best of all, implying a potential benefit of *both* rehearsal and retrieval in that condition. However, the fact that the Rehearsal-only condition was worse than the Neither condition (and getting worse with training) suggests that rehearsal was of little help; the Both and Retrieval-only conditions may have continued to perform equivalently, even after the potential ceiling effect was eliminated.

The external cue presented in the current experiments was very explicit. In fact, it reiterated the original prospective command (GO TO X). As such, the external cue was not strictly speaking a memory cue in the Both and Retrieval-only conditions. If the subjects remembered to look at the proper strip (assuming that they could not remember the information unassisted), the command was there. We could have presented the command initially, and then manipulated the temporal availability of a less explicit cue (e.g., the word CHANGE rather than GO TO X). Subjects would not perform as accurately in this modified experiment. However, we would expect the same pattern of results (Both = Retrieval-only > Rehearsal-only = Neither) if we scored the data as follows: *Correct* if plane *not* sent to its original data block destination, *Wrong* if plane sent to original data block destination. That is, *Correct* means that subjects knew to make a change (the data block destination was wrong). Having no idea what the correct destination was, however, they could only guess among the five remaining destinations.

How general is the conclusion that external cues support primarily a retrieval function? What circumstances might produce a different result? We consider several mitigating factors which, in a different situation, could result in different conclusions being drawn.

The relationship between the Both and Retrieval-only conditions may depend on the distinctiveness of the cue (McDaniel & Einstein, 1993, showed that distinctive cues were better than nondistinctive cues). For example, the success of the Both condition may depend on whether subjects habituate to a continuously available cue. Stepping over a to-be-returned library book throughout the course of the day may increase the likelihood that you will simply step over it yet again on your way home. Conversely, the success

of the Retrieval-only condition may depend on the attention-grabbing ability of the external cues when it reappears. In the present experiments, we decided to *not* make the return of the external cue conspicuous. Clearly, a Retrieval-only cue could be designed such that the cue would never be missed (reappear in bright orange at twice its normal size). Such a cue could serve as a trigger, as well as support the content of the action. We will have more to say about this when we discuss the implications of this research for interface design.

The success of the Rehearsal-only condition likely depends on the type of rehearsal used during the retention interval. Maintenance rehearsal does not typically aid retrospective performance (Craik & Watkins, 1973), but elaborative rehearsal might, at least it typically does in retrospective memory tasks (e.g., Craik & Tulving, 1975). Subjects had the opportunity to tailor their rehearsal to be maximally effective in Experiment 2, and performance did not improve relative to the Neither condition. The ability to make use of elaborative rehearsal may be workload dependent; performance on the air traffic control simulation may have begun to suffer if subjects tried more elaborative rehearsal strategies. Indeed, the task was fairly difficult, although it became easier over days. If this improvement was accompanied by a decrease in workload, we might have expected to see the Rehearsal-only condition improve with training. In fact, the opposite occurred.

The results of Vortac et al. (1993) implied that automation may benefit prospective memory. The present results are a first look at how external cues might be presented to take advantage of the automation. However, which is to be preferred: having the external cue present throughout the retention interval, or only when it is needed? What additional factors might lead an interface designer to choose one method of presenting cues over the other?

The final design of a human-computer interface involves many tradeoffs (see Rubin, 1988). For example, if density of information is a problem, it may be preferable to implement the external cue like the Retrieval-only condition: the cue will clutter the screen for less time. In addition, a late-returning external cue in the Retrieval-only condition could be made to be novel and therefore grab attention. An external cue that grabs attention (or sounds a warning) might seem like a good feature, but not if the user's attention is taken away from a more important task, or taken away at an inopportune time. On the other hand, if workload is a limiting factor, the extra time spent attaching a time tag to the Retrieval-only condition might be too costly and it might be better to simply make the cue available for the whole time. However, the user then runs the risk of habituating to the constantly present cue.

The present research succeeded in disentangling two confounded functions of external cues, rehearsal and retrieval. We used a task of obvious relevance to air traffic controllers, although we suspect that the conclusions hold for anyone who works in a highly complex dynamic environment that puts a strain on working memory. Our everyday cognitive functioning is just such an environment.

References

Aerospace Sciences, Inc. (1991, November). *Air traffic control specialist pre-training screen preliminary validation: Final report.* (FAA Report No. DTFA-01-90-Y-01034). Fairfax, VA.

Baddeley, A. D., & Wilkins, A. J. (1984). Taking memory out of the laboratory. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory, actions, and absentmindedness* (pp. 1-17). New York: Academic Press.

Ceci, S. J., & Bronfenbrenner, U. (1985). "Don't forget to take the cupcakes out of the oven": Prospective memory, strategic time-monitoring, and context. *Child Development, 56*, 152-164.

Cox, M. (1992). *The cognitive aspects of the air traffic control task: A literature review.* Farnborough, Hampshire, UK: Royal Air Force Institute of Aviation Medicine, Report No. 718).

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology, 104*, 268-294.

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 12*, 599-607.

Drew, G. C. (1940). An experimental study of mental fatigue. *British Air Ministry, Flying Personnel Research Committee Paper No. 277.*

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 717-726.

Harris, J. E. (1984). Remembering to do things: A forgotten topic. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory, actions, and absent-mindedness.* London, England: Academic Press.

Hopkin, V. D. (1988). *Human factors aspects of the AERA 2 program.* Farnborough, Hampshire, UK: Royal Air Force Institute of Aviation Medicine.

Intons-Peterson, M. J., & Fournier, J. (1986). External and internal memory aids: How often do we use them? *Journal of Experimental Psychology: General, 115,* 267-280.

Jackson, A. (1989). *The functionality of flight strips.* (Human Factors Working Note). Gt. Malvern, UK: Royal Signal and Radar Establishment.

McDaniel, M. A., & Einstein, G. O. (1993). The importance of cue familiarity and cue distinctiveness in prospective memory. *Memory, 1,* 23-41.

Meacham, J. A. & Colombo, J. A. (1980). External retrieval cues facilitate prospective remembering in children. *Journal of Educational Research, 73,* 299-301.

Meacham, J. A. & Leiman, B. (1982). Remembering to perform future actions. In U. Neisser (Ed.), *Memory observed: Remembering in natural context* (pp. 327-336). San Francisco: Freeman.

Meacham, J. A. & Singer, J. (1977). Incentive effects in prospective remembering. *The Journal of Psychology, 97,* 191-197.

Petro, S. J., Herrmann, D., Burrows, D., & Moore, C. M. (1991). Usefulness of commercial memory aids as a function of age. *International Journal of Aging and Human Development, 33,* 295-309.

Rubin, T. (1988). *User interface design for computer systems.* Chichester, West Sussex, England: Ellis Horwood Limited.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592-604.

Vortac, O. U. (1993). Should Hal open the pod bay doors? An argument for modular automation. In D. J. Garland & J. A. Wise (Eds.). *Human factors and advanced aviation technologies.* Daytona Beach, Florida: Embry-Riddle Aeronautical University Press.

Vortac, O. U., Edwards, M. B., Fuller, D. K., & Manning, C. A. (1993). Automation and cognition in air traffic control: An empirical investigation. *Applied Cognitive Psychology, 7,* 631-651.

Vortac, O. U., & Gettys, C. F. (1990). *Cognitive factors in the use of flight progress strips: Implications for automation.* Cognitive Processes Laboratory Working Paper, University of Oklahoma, Norman, OK.

West, R. (1988). Prospective memory and aging. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues. Volume 2: Clinical and educational implications.* (pp. 119-125). Chichester, England: Wiley.

Wilkins, A. J., & Baddeley, A. D. (1978). Remembering to recall in everyday life: An approach to absent-mindedness. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.) *Practical aspects of memory.* London, England: Academic Press.

☆U.S. GOVERNMENT PRINTING OFFICE: 1995 - 660-067/00093